What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

February 2013

WWC Review of the Report "Summer School Effects in a Randomized Field Trial"[1]

The findings from this review do not reflect the full body of research evidence on summer literacy interventions.

What is this study about?

The study examined the impact of a summer literacy program on kindergarten and first-grade students who were at moderate risk for reading difficulties in one Pacific Northwest school district.

The study took place through a limited expansion of an existing summer program for high-risk students that was modified to include moderate-risk students. Study authors randomly assigned 49 kindergarten students (25 intervention, 24 comparison) and 51 first-grade students (26 intervention, 25 comparison) identified as moderate-risk to either an intervention group that was invited to participate in the summer reading program, or a comparison group that did not receive the intervention.[2] The final analytic sample consisted of 46 kindergarten students (24 intervention, 22 comparison) and 47 first-grade students (23 intervention, 24 comparison).

The study assessed the effectiveness of the summer literacy program by comparing the achievement of the intervention and comparison students. For students in kindergarten, the outcome of interest was an alphabetic assessment. For students in first grade, the outcome of interest was student reading fluency.

WWC Rating

The research described in this report meets WWC evidence standards without reservations

Strengths: This study was a well-implemented randomized controlled trial with low attrition.

Features of the summer literacy intervention

The 5-week district-sponsored literacy program was designed for early-elementary students during the summer following kindergarten or first grade. The program was held 4 days per week, for 3.5 hours a day, in the middle of the 3-month summer break. Instruction was aligned with "big ideas" and best practices from the National Reading Panel (2000). Each class consisted of 2 hours of teacher-directed instruction in phonemic awareness, alphabetic understanding, and fluency/automaticity. Led by a teacher, students then practiced the skills in groups of three to five.

What did the study find?

The study found, and the WWC confirmed, a statistically significant positive effect of the summer school intervention on student outcomes in the fall of the implementation year for students in both kindergarten (effect size on the alphabetic assessment = 0.69) and first grade (effect size on the reading fluency assessment = 0.61).

Appendix A: Study details

Zvoch, K., & Stevens, J. J. (2012). Summer school effects in a randomized field trial. *Early Childhood Research Quarterly, 28*(1), 24–32. Retrieved from http://dx.doi.org/10.1016/j.ecresq.2012.05.002

Setting

The study was conducted in one school district located in a moderately sized city in the Pacific Northwest. At the time of the study, the district had a summer literacy intervention in place for students who were at high risk for reading difficulties at the end of kindergarten or first grade. Risk levels were determined by students' performing under designated thresholds on an end-of-grade assessment covering alphabetics for kindergarten students and reading fluency for first-grade students. This evaluation took place through a limited expansion of the summer program that allowed students at moderate risk of reading difficulties to be eligible to attend the program.

Study sample

After 49 students in kindergarten and 51 students in first grade were identified as being at moderate risk for reading difficulties in the spring, approximately half were randomly assigned the opportunity to attend the summer literacy program. Although 25 students in kindergarten and 26 students in first grade were offered the program, only 13 students in kindergarten and 17 students in first grade actually participated in the summer program. However, the authors of the study performed an intent-to-treat analysis, comparing the achievement levels of the students who were *offered* the opportunity to attend the summer literacy institute with the students who were *not offered* this opportunity. The final analysis sample included 24 students in kindergarten and 23 students in first grade who were offered the intervention, and 22 students in kindergarten and 24 students in first grade who were not offered the intervention.

Intervention group

The intervention was a district-sponsored, 5-week literacy program for early elementary students. It was implemented for 3.5 hours a day, 4 days a week, in the middle of the 3-month summer break. Instruction was aligned with the "big ideas" and best practices from the National Reading Panel (2000). Each class consisted of 2 hours of teacher-directed instruction in phonemic awareness, alphabetic understanding, and fluency/automaticity. This was followed by practice in groups of three to five students. Each class consisted of approximately 20 students and four teachers.

Comparison group

The students in the comparison condition were not offered, and therefore did not attend, the summer literacy intervention. This condition was described as a "no-treatment" comparison group.

Outcomes and measurement

The study assessed students in both grades in both the spring (pretest) and the fall (posttest). Students in kindergarten were assessed using the Nonsense Word Fluency (NWF) subtest of the Dynamic Indicators of Basic Early Literacy Skills (DIBELS), which is an outcome in the alphabetics domain. Students in first grade were assessed using the Test of Oral Reading Fluency (TORF), which is an outcome in the reading fluency domain.

Support for implementation

No information about support for implementation was provided.

Reason for review

This study was identified for review by the WWC because it was supported by a grant to the University of Oregon (Principal Investigator: Keith Zvoch) from the National Center for Education Research (NCER) at the Institute of Education Sciences (IES).

Appendix B: Outcome measures for each domain

Alphabetics	
Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Nonsense Word Fluency (NWF) subtest	The DIBELS NWF subtest is a standardized assessment that was administered to the students in the kindergarten sample. Each student is presented with a random series of nonsense words (e.g., vad, ab) and is tasked with orally reproducing the sound of the target word. A student's score is determined by the number of correct letters they can correctly reproduce from the series of nonsense words in one minute. Test-retest reliability of the assessment, based on alternate-form reliability from a different student sample, was 0.83.
Reading fluency	
Test of Oral Reading Fluency (TORF)	The TORF is a standardized assessment that was administered to the students in the first-grade sample. Each student is tasked with reading three passages aloud for one minute each. When a word is omitted, substituted, or the student hesitates for more than three seconds, it is scored as an error. Any word that is self-corrected within a three-second timeframe is scored as accurate. The median words-per-minute across the three passages is used as the student's score. Test-retest reliability was reported by the test developer to range from 0.92 to 0.97.

Appendix C: Study findings for each domain

Domain and outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Alphabetics								
DIBELS NWF subtest	Kindergarten students	46 students	39.11 (13.59)	30.27 (11.58)	8.84	0.69	+25	< 0.05
Reading fluency								
TORF	First-grade students	47 students	55.24 (14.04)	46.04 (15.34)	9.20	0.61	+23	< 0.05

Table Notes: Positive results for mean difference, effect size, and improvement index favor the intervention group; negative results favor the comparison group. The effect size is a standardized measure of the effect of an intervention on student outcomes, representing the change (measured in standard deviations) in an average student's outcome that can be expected if the student is given the intervention. The improvement index is an alternate presentation of the effect size, reflecting the change in an average student's percentile rank that can be expected if the student is given the intervention. The effect reported in each of the domains is positive and statistically significant.

Study Notes: The mean difference, effect size, and improvement index are based on WWC calculations using the unadjusted pretest and posttest means presented in the study. The study presented achievement results separately for those receiving the intervention, those who were offered but refused the intervention, and those who were not offered the intervention. Based on an intent-to-treat analysis, the WWC calculations first combined those receiving the intervention and those who refused. The WWC calculated the intervention group mean by adding the difference-in-differences adjusted estimate of the average impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. Please see the WWC Handbook for more information. No corrections for clustering or multiple comparisons were needed, and the *p*-values presented here were reported in the original study.

Endnotes

[1] Single study reviews examine evidence published in a study (supplemented, if necessary, by information obtained directly from the author[s]) to assess whether the study design meets WWC evidence standards. The review reports the WWC's assessment of whether the study meets WWC evidence standards and summarizes the study findings following WWC conventions for reporting evidence on effectiveness. This study was reviewed using the Beginning Reading review protocol, version 2.0.

[2] The numbers of students initially assigned to the intervention and comparison conditions were obtained from the author in email correspondence.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, February). *WWC review of the report: Summer school effects in a randomized trial*. Retrieved from http://whatworks.ed.gov.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Improvement index Along a percentile distribution of students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which subjects are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which investigators randomly assign eligible participants into intervention and comparison groups.

Single-case design (SCD) A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample are spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < 0.05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Please see the WWC Procedures and Standards Handbook (version 2.1) for additional details.